THE JOURNAL OF SYMBOLIC LOGIC
Volume 6, Number 1, March 1941

A SYSTEM OF AXIOMATIC SET THEORY—PART II[13]

PAUL BERNAYS

4. The axioms, second part. For the formulation of the remaining axioms we need the notions of a *function* and of a *one-to-one correspondence*.

We define a *function* to be a class of pairs in which different elements always have different first members; or, in other words, a class F of pairs such that, to every element a of its domain there is a unique element b of its converse domain determined by the condition $\langle a, b\rangle \eta F$. We shall call the set b so determined the *value* of F for a, and denote it (following the mathematical usage) by $F(a)$.

A set which represents a function—i.e., a set of pairs in which different elements always have different first members—will be called a *functional set*.

If b is the value of the function F for a, we shall say that F *assigns* the set b to the set a; and if a functional set f represents F, we shall say also that f *assigns* the set b to the set a.

A class of pairs will be called a *one-to-one correspondence* if both it and its converse class are functions. We shall say that there exists a one-to-one correspondence between the classes A and B (or of A to B) if A and B are domain and converse domain of a one-to-one correspondence. Likewise we shall say that there exists a one-to-one correspondence between the sets a and b (or of a to b) if a and b respectively *represent* the domain and the converse domain of a one-to-one correspondence. In the same fashion we speak of a one-to-one correspondence between a class and a set, or a set and a class.

IV. AXIOM OF CHOICE. Every class C of pairs has a subclass which is a function and has the same domain as C.

Consequences.

1. For every class C of non-empty sets there exists a function, having C as its domain, which assigns to every set belonging to C one of its elements.—For, by the axioms III, the class of pairs $\langle a, b\rangle$ such that $a \eta C$ and $b \epsilon a$ exists; and we apply the axiom of choice to this class.

2. For every function F there exists an *inverse function*, i.e., a function which is a subclass of the converse class of F and whose domain is the converse domain of F.—This follows immediately by applying the axiom IV to the converse class of F.

Remark. If the consequence 2 (just stated) is taken as an axiom, then the axiom IV can be inferred from it. In fact, by the class theorem, for every class C of pairs there exists the class of all pairs of the form $\langle\langle a, b\rangle, a\rangle$ where $\langle a, b\rangle \eta C$. This class is a function whose converse domain is the domain of C. If for every function there is an inverse function, we may infer the existence of a function,

Received June 12, 1940.

[13] Part I appeared in this JOURNAL, vol. 2 (1937), pp. 65–77.

having the same domain as C, whose elements are of the form $\langle a, \langle a, b\rangle\rangle$ where $\langle a, b\rangle \eta C$. The converse domain of this function is a function having the same domain as C and is a subclass of C.

Thus the axiom IV is equivalent in content to the assumption that for every function there exists an inverse function.

V. Axioms concerning the representation of classes by sets.[14]

a. If a class is represented by a set, every subclass of it is also represented by a set. (Subclass axiom.)

b. If the domain of a one-to-one correspondence is represented by a set, the converse domain is also represented by a set. (Axiom of replacement.)

c. For every set a, the class which is the sum of the elements of a is represented by a set. (Sum axiom.)

d. For every set a, the class of all subsets of a is represented by a set. (Power axiom.)

Consequences (1–9). Since, by consequence 7 of the axioms III, every set represents a class, we have as an immediate consequence of V a:

1. Every subclass of a set is represented by a set.

As a consequence of 1 we have:

2. The intersection of the elements of a non-empty class or of a non-empty set is represented by a set.—For, if a is an element of a class C (or of a set c), the intersection of the elements of C (or of c) is a subclass of a.

From V b we infer:

3. If the domain of a function is represented by a set, the function itself is represented by a functional set; and conversely, the domain of a function represented by a functional set is represented by a set.—For, by the class theorem, if F is a function, the class of pairs $\langle a, \langle a, b\rangle\rangle$ such that $\langle a, b\rangle \eta F$ exists; this class is a one-to-one correspondence, its domain is the domain of F, and its converse domain is the class F itself.

As a particular application of 3, we note the following. The condition that there exists a one-to-one correspondence between the sets a and b is equivalent to the condition that there is a set representing a one-to-one correspondence between a and b. The latter condition can be symbolically formulated by a constitutive expression, since it amounts to the condition that there is a set s of pairs such that (1) different elements of s always differ in both their first and their second members, and (2) the elements of a are just those sets which occur as first members of elements of s and the elements of b are just those sets which occur as second members of elements of s. Hence, by the class theorem, there exists the class of all pairs $\langle a, b\rangle$ for which a one-to-one correspondence between a and b exists.

[14] The axioms V a, c, d correspond respectively to Zermelo's *Axiom der Aussonderung*, *Axiom der Vereinigung*, and *Axiom der Potenzmenge*. Cf. E. Zermelo, *Untersuchungen über die Grundlagen der Mengenlehre I*, ***Mathematische Annalen***, vol. 65 (1908), pp. 261–281. Axiom V b is a modification of Fraenkel's *Axiom der Ersetzung*. Cf. A. Fraenkel, *Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre*, ***Mathematische Annalen***, vol. 86 (1922), pp. 230–237; and ***Zehn Vorlesungen über die Grundlegung der Mengenlehre***, Leipzig and Berlin 1927, see p. 115 and pp. 104–110.

From IV, V a, and V b follows:

4. If the domain of a function is represented by a set, the converse domain is also represented by a set.—For let F be a function whose domain A is represented by a set, and let B be the converse domain of F. By consequence 2 of the axiom IV, there is an inverse function G to F, whose domain is B, and which is a subclass of the converse class C of F. The converse class of G is a one-to-one correspondence between a subclass A_1 of A, and B. Since A is represented by a set, it follows from V a that A_1 is represented by a set, and hence, by V b, that B is represented by a set.

Remark. The consequence 4, which we have just derived from the axioms IV, V a, V b, will be called the *theorem of replacement.* It has about the same content as Fraenkel's axiom of replacement in its original form.

If 4 is taken as an axiom, V a and V b become provable as theorems. This is obvious in the case of V b. To prove V a we must show that a subclass S of a class A represented by a set is represented by a set. This is immediate if S is empty. Otherwise let c be an element of S. By the class theorem, there exists the class of those pairs $\langle a, b\rangle$ for which, either $b=a$ and $a\eta S$, or $b=c$ and $a\eta A$ but not $a\eta S$. This class is a function whose domain is A and whose converse domain is S. Since A is represented by a set, it follows from 4 (taken as an axiom) that S is represented by a set.

Thus we have the possibility of replacing the axioms V a, b by the theorem of replacement taken as an axiom (which we shall refer to as axiom V*). This modification is appropriate in connections where the intention is to get along as far as possible without the axiom of choice.

A further remarkable possibility is that of taking instead of V a, b, c the following axiom, which we shall call V**: If the domain of a function is represented by a set, then the sum of the elements of the converse domain (i.e., the sum of the values of the function) is represented by a set.

In fact, from V** we can obtain V c and the theorem of replacement, from which latter, as we have just seen, V a and V b can be derived. If a is any set, there exists, by the class theorem, the class of those pairs $\langle c, c\rangle$ for which $c\epsilon a$. Since this class is a function whose domain is a and whose converse domain is a, we conclude that the sum of the elements of a is represented by a set. Moreover, by the class theorem, if F is any function, there exists the class of those pairs $\langle c, d\rangle$ for which c belongs to the domain of F and $d=(F(c))$. Since this class is a function, and the sum of the values of this function is the class of the values of F—i.e., the converse domain of F—we conclude that if the domain of F is represented by a set the converse domain is also represented by a set.

On the other hand V** follows directly from the theorem of replacement and V c; hence it may be derived as a theorem from IV, V a, V b, V c. We shall call this theorem—whose statement coincides with that of the axiom V**—the *sum theorem.*

As a consequence of 3 and 4 above, we have:

5. If a function is represented by a functional set, its domain and its converse domain are each represented by a set.

Definition. If f is a functional set representing a function F, we shall call the set representing the domain of F the *domain of f*, and the set representing the converse domain of F the *converse domain of f*. Thus the domain and the converse domain of a functional set are, by definition, sets.

From IV and V b, there follows:

6. If s is a set of which 0 is not an element, and if no two elements of s have a common element, then there exists a set c, such that every element of c is in some element of s, and c has one and only one element in common with each element of s.[15]—For let S be the class represented by s (consequence 7 of the axioms III). Then, by consequence 1 of the axiom IV, there exists a function F, with S as its domain, which assigns to every element a of S one of the elements of a. Since no two elements of S have a common element, F is a one-to-one correspondence. Hence, by the axiom of replacement, the converse domain of F is represented by a set c; and it is easily seen that this set c has the required properties.

As we thus see, 6 results from combining a statement about the existence of a function (consequence 1 of our axiom IV) with the axiom of replacement. This was pointed out by Fraenkel.[16]

From V c we obtain:

7. If the classes A and B are represented by sets, their sum is represented by a set.—For if A and B are represented by a and b, the sum of A and B is the sum of the elements of the set (a, b).

From V b, c we infer:

8. If the classes A and B are represented by sets, the class of those pairs $\langle a, b\rangle$ for which $a \eta A$ and $b \eta B$ is represented by a set.—For, by the axioms III, corresponding to every fixed element a of A there exists the class of those pairs $\langle b, \langle a, b\rangle\rangle$ for which $b \eta B$. Since this class is a one-to-one correspondence and its domain B is represented by a set, the converse domain is, by V b, also represented by a set. Thus for every element a of A there exists the *set* of those pairs $\langle a, b\rangle$ for which $b \eta B$. Moreover, by the class theorem, there exists the class C of those pairs $\langle a, c\rangle$ whose first member a belongs to A, and whose second member c is the set of all pairs having a as first member and an element of B as second member. It follows that the domain of C is A. Hence, by the axiom of replacement, since C is a one-to-one correspondence and A is represented by a set, the converse domain of C is represented by a set. Hence, by V c, the sum

[15] This is the assertion of the *multiplicative axiom*, which was first introduced by Russell (in his paper *On some difficulties in the theory of transfinite numbers and order types*, ***Proceedings of the London Mathematical Society***, ser. 2 vol. 4 (1906–7), see pp. 47–52) as a modification of the original Zermelo choice postulate (cf. *Beweis, dass jede Menge wohlgeordnet werden kann*, ***Mathematische Annalen***, vol. 59 (1904), see p. 516), and was independently stated by Zermelo as *axiom of choice*, from which he was able to infer, within his axiom system, the assertion of his former postulate (cf. *Neuer Beweis für die Möglichkeit einer Wohlordnung*, ***Mathematische Annalen***, vol. 65 (1908), see p. 110, and *Untersuchungen über die Grundlagen der Mengenlehre I*, ***Mathematische Annalen***, vol. 65 (1908), see p. 266 and pp. 273–274).

[16] *Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre*, ***Mathematische Annalen***, vol. 86 (1922), pp. 230–237, see p. 233.

of the elements of the converse domain of C is represented by a set. But this sum is just the class of those pairs $\langle a, b\rangle$ for which $a\eta A$ and $b\eta B$.

We can also derive 8 from V a, c, d in the following way. If the classes A and B are represented by sets, it follows from 7 (which is a consequence of V c) that the sum of A and B is represented by a set s. By V d, the class of subsets of s is represented by a set t, and the class of subsets of t by a set u. Now every pair $\langle a, b\rangle$ for which $a\eta A$ and $b\eta B$ is an element of u; for we have $a\epsilon s$ and $b\epsilon s$, hence $(a)\epsilon t$ and $(a, b)\epsilon t$, hence $((a), (a, b))\epsilon u$, i.e., $\langle a, b\rangle\epsilon u$ —since, by definition, $\langle a, b\rangle = ((a), (a, b))$. Consequently the class of pairs whose first members belong to A and second members to B is a subclass of u, and so, by V a, is represented by a set.

This method of proving 8 is more familiar to most mathematicians than that by V b, c. There is also another method of proving 8, namely by means of IV, V a, V b without using V c or V d, which we shall indicate later on.

Finally, from V a, c, d we obtain the consequence:

9. If a and b are sets, there exists a set whose elements are those functional sets having a as domain and a subset of b as converse domain.—For, since every set represents a class, it follows from 8 that, given sets a and b, there exists a set c whose elements are those pairs $\langle k, l\rangle$ for which $k\epsilon a$ and $l\epsilon b$. Every functional set whose domain is a and whose converse domain is a subset of b is a subset of c; hence the class M of all such functional sets—which exists by the class theorem—is a subclass of the class D of all subsets of c. Now, by V d, D is represented by a set. Hence, by V a, M is represented by a set m, which is the set whose existence is asserted in 9.

VI. Axiom of infinity. There exists a set of which there is a one-to-one correspondence to a proper subclass. Or, in other words, there exists a one-to-one correspondence of a class to a proper subclass, the domain of the one-to-one correspondence being represented by a set.

From this axiom, with either V a or V b, we get the *immediate consequence*: There exists a set of which there is a one-to-one correspondence to a proper subset. This assures the existence of a set representing an infinite class, under the Dedekind definition of infinity,[17] according to which a class is called infinite if there exists a one-to-one correspondence between it and a proper subclass of it.

We shall later discuss the question of the equivalence of this axiom VI to other forms of an axiom of infinity, i.e., of an axiom guaranteeing the existence of an infinite set.

Remark. The axiom of infinity in *Principia mathematica* is of quite a different kind from the axioms of infinity with which we are here concerned. For it deals with the existence of infinitely many individuals—which, in our system, is a consequence of the axioms II.

[17] Cf. Richard Dedekind, ***Was sind und was sollen die Zahlen?***, Braunschweig 1888, §5. —Of course Dedekind makes no distinction between *classes* and *sets*; as a matter of fact he uses neither word, but speaks of *systems*.

VII. Restrictive axiom.[18] To every non-empty class A there belongs a set b such that there is no common element of A and b. Or, in other words, there is no non-empty class A every element of which has an element belonging to A.

Since every set represents a class, an immediate consequence of VII is:

1. Every non-empty set a has an element b such that a and b have no common element.

On the other hand, from this assertion taken as an axiom (call it axiom VII*) axiom VII can be derived as a theorem—as we shall show later—either[19] by means of V a, b, c, d and VI or by means of IV, V a, V b, VI.

From 1 we infer:

2. There is no reflexive set.—For if $a \epsilon a$, the set (a) would contradict 1.

In an analogous way there follows:

3. There are no sets a, b such that $a \epsilon b$ and $b \epsilon a$. There are no sets a, b, c such that $a \epsilon b$ and $b \epsilon c$ and $c \epsilon a$.

4. The null set is an element of every non-empty transitive set.—For let a be a non-empty transitive set. By 1, a has an element c which has no element that is an element of a. But, a being transitive, every element of c is an element a. Therefore c is the null set.

5. Foundations of the theory of ordinal numbers. In axiomatic set theory, the ordinal numbers can be introduced without referring in the definition to the concept of order, so that the theory of ordinals is obtained independently of the theory of ordered sets.

Such an independent general theory of ordinals was obtained by Zermelo (about 1915) but has not been published.

A rather convenient form of this independent theory of ordinals has recently been exhibited by Raphael M. Robinson in connection with his modification of the von Neumann system.[20] We shall follow his method, the simplicity of which will perhaps appear still more clearly in the adaptation to our system.

The only axiom required for this purpose in addition to the elementary axioms I–III is the restrictive axiom VII. As we shall see later, the necessity for axiom VII can be avoided by a modification in the definition of an ordinal, but in that case the axiom V a becomes necessary.

Definition. An *ordinal* (or an *ordinal number*) is a transitive set such that, if a and b are any two different elements of it, either $a \epsilon b$ or $b \epsilon a$.

---

[18] The idea of this axiom was first conceived by von Neumann in his paper *Eine Axiomatisierung der Mengenlehre* (***Journal für die reine und angewandte Mathematik,*** vol. 154 (1925), see p. 239, axiom VI 4). From the original form of it he passed to the present form of the axiom, or to a somewhat stronger axiom in a form adapted to his system, in *Über eine Widerspruchsfreiheitsfrage in der axiomatischen Mengenlehre* (***Journal für die reine und angewandte Mathematik,*** vol. 160 (1929), see p. 231). Zermelo independently introduced this axiom, calling it *Axiom der Fundierung*, in his paper *Über Grenzzahlen und Mengenbereiche* (***Fundamenta mathematicae,*** vol. 16 (1930), see p. 31).

[19] This dependency was stated by Gödel.

[20] *The theory of classes, a modification of von Neumann's system*, this Journal, vol. 2 (1937), pp. 29–36.

In consequence of this definition, by the class theorem, the *class of all ordinals* exists.

For the development of the elementary theory of ordinals, we require the following lemmas about reflexive and transitive sets, proofs of which are obvious.

Lemma 1. Every reflexive set has at least one reflexive element.

Lemma 2. If all the elements of a non-empty class are transitive, the set representing the intersection of the elements of the class is also transitive.

Lemma 3. If all the elements of a non-empty class are transitive, a set representing the sum of the elements of the class must also be transitive.

Definition. By consequence 2 of the axioms I, II, there is, corresponding to any set c, a unique set whose elements are c itself and the elements of c. This set will be called c'.

Lemma 4. If a is a transitive set, a' is also transitive.

Let us extend the notion of transitivity to classes as well as sets, calling a class transitive if every element of an element of it is an element of it—or, in other words, if every element of it is a subset of it. Then, corresponding to lemma 2 we have: The intersection of the elements of a class of transitive sets is a transitive class.

We now go on to the fundamental theorems on ordinals.

Theorem 1. Every transitive proper subclass of an ordinal is represented by an element of the ordinal.

For let n be an ordinal and M a transitive proper subclass of n. If then D is the class of those elements of n not belonging to M, D is not empty. By VII, there exists an element c of D such that there is no common element of c and D. Now n (as an ordinal) is transitive. Hence every element of c is in n and, since it cannot belong to D, must belong to M. Therefore c is a subset of M.

On the other hand, since $c \eta D$, c does not belong to M; and, because M is transitive, c cannot be in an element of M. Thus if $a \eta M$, neither $a = c$ nor $c \epsilon a$. But $a \epsilon n$ (since M is a subclass of n), and n is an ordinal. Therefore $a \epsilon c$; and, since a was an arbitrary element of M, it follows that M is a subclass of c.

Since M is a subclass of c and c is a subset of M, we have the result that c represents M.

As a corollary of this theorem we have: Every transitive proper subset of an ordinal is an element of it. In particular, if a and b are ordinals, and $a \subset b$, then $a \epsilon b$.

Theorem 2. The intersection of the elements of a non-empty class of ordinals is represented by a set belonging to this class.

For if C is a non-empty class of ordinals, and D is the intersection of the elements of C, then, by the generalization of lemma 2, D is a transitive class. Let a be any element of C; then D is a subclass of a. If D is represented by a, theorem 2 is verified. Otherwise D is a transitive proper subclass of a and consequently, by theorem 1, is represented by a set d which is an element of a. However, D must be represented by *some* element of C, because otherwise d would be an element of every element of C, and from this would follow $d \eta D$, and therefore $d \epsilon d$, in contradiction to consequence 2 of axiom VII.

Consequences. Since every set represents a class, it follows that a set, representing the intersection of the elements of a non-empty set c whose elements are ordinals, is itself in c. This applies in particular to the case that c is a set (a, b), a and b being ordinals. Hence follows that the intersection of two ordinals is one of them. I.e., if a and b are distinct ordinals, either $a \subset b$ or $b \subset a$.

Thus the ordinals are *ordered* by the relation "$\subset$" ("being a proper subset of"). Of two ordinals, the one which is a proper subset of the other will be called the *lower*, and the other one, the *higher*. This ordering of the ordinals is a *well-ordering*; in fact it follows from theorem 2 that *among the elements of a non-empty class of ordinals there is always a lowest one.*

Notice that we are here using the words *ordering* and *well-ordering* in the ordinary mathematical sense. Definitions of them within our system, which are here not yet required, will be given later on.

Since, if a and b are ordinals, the relation $a \subset b$ entails $a \epsilon b$, we have that, of any two different ordinals, the lower one is an element of the higher. Thus if a and b are distinct ordinals, either $a \epsilon b$ or $b \epsilon a$.

From this there follows further that *every transitive set of ordinals is an ordinal.*

Theorem 3. Every element of an ordinal is an ordinal.

For let n be an ordinal and c an element of it. Since n is transitive, c is a subset of n. Therefore if a and b are any two different elements of c, either $a \epsilon b$ or $b \epsilon a$. But c is also transitive; for if $a \epsilon b$ and $b \epsilon c$, the possibilities $a = c$ and $c \epsilon a$ are excluded by consequence 3 of axiom VII; therefore, since a and c are both elements of the ordinal n, we must have $a \epsilon c$. Thus c is an ordinal.

Combining with theorem 3 our results concerning the ordering of ordinals, we obtain: *Every ordinal n represents the class of ordinals lower than n.*

Since every transitive set of ordinals is an ordinal, we have further, by theorem 3 and lemma 4: *If a is an ordinal, a' is also an ordinal.*

In a similar way, using lemma 3 instead of lemma 4, we obtain: If a set represents the sum of the elements of a class of ordinals, it is itself an ordinal.

The following results are also easily obtained. 0 is the lowest ordinal. If a is an ordinal, the next higher ordinal is a'. If a and b are ordinals, and $a' = b'$, then $a = b$. An ordinal a' has a as its highest element; and conversely, an ordinal having a highest element a must be a'. If C is a class of ordinals which has no highest element, and there is an ordinal which is higher than every element of C, then the sum of the elements of C is represented by a set, namely by the lowest of those ordinals which are higher than every element of C.

We have also the negative result that *the class of all ordinals cannot be represented by a set.* In fact, such a set would be a transitive set of ordinals, by theorem 3, and would consequently be itself an ordinal; but this would make it reflexive, contrary to consequence 2 of axiom VII.

Now it is readily verified that the applications of axiom VII which we have made in connection with the theory of ordinals all refer to subclasses of ordinals. In consequence, as remarked by R. M. Robinson,[21] it would be possible to avoid the use of this axiom by adjoining to the definition of an ordinal the following

[21] Loc. cit. (footnote 20).

additional condition which a set must satisfy in order to be an ordinal: To every non-empty subclass C of the set there belongs a set which has no element belonging to C.

But this modified definition does not allow us to infer by means of the class theorem that the class of all ordinals exists, although this result is necessary for our purposes. The difficulty can be removed, however, if we allow the use of V a; for from this axiom it follows that the stated condition on a subclass C is equivalent to the corresponding condition on a subset (cf. consequence 1 of V a).

Indeed, on the basis of the axioms I–III and V a, we can derive the theory of ordinals in essentially the same way as before if we replace our former definition of an ordinal by the following alternative or second definition:

A set n is an ordinal if (1) n is transitive, and (2) for any two different elements a and b of n either $a\epsilon b$ or $b\epsilon a$, and (3) every non-empty subset s of n has an element c such that c and s have no common element.

Only the two following points then have to be taken into account:

1. In order to conclude that every transitive set of ordinals is itself an ordinal, we must now also show that such a set satisfies the condition (3) of our (second) definition of an ordinal; but this follows at once from the fact that in every non-empty set of ordinals there is a lowest ordinal.

2. In the proof of theorem 3 we have to show that every element of an ordinal satisfies the condition (3); but this is immediate, since an ordinal is transitive and so every element of it is a subset of it.

We add the following remarks:

1) The equivalence of our second definition of an ordinal to the first definition, which obviously subsists on the basis of axiom VII, ceases to be provable with V a in the place of VII. For without VII we cannot exclude, e.g., the possibility of a reflexive set which is its own only element—as we shall show later on by means of a model. Such a set would be an ordinal under our first definition, but not under the second.

We shall usually understand the word *ordinal* in the sense of our second definition—but this can be replaced by the simpler first definition in the case that axiom VII is assumed.

2) With the aid of axiom VII we can prove that a transitive set all of whose elements are transitive is an ordinal. In fact, let n be a transitive set all of whose elements are transitive, and let S be the class of those elements of n which are not ordinals. If S were not empty, there would follow from axiom VII the existence of an element c of S having no element belonging to S. Since n is transitive, every element of c would be an element of n not belonging to S and consequently an ordinal. Moreover, as an element of n, c would be transitive. But a transitive set of ordinals is itself an ordinal; so c would be an ordinal and could not belong to S. The class S must therefore be empty. Thus n is a transitive set of ordinals and is therefore itself an ordinal.

The converse of this theorem is an immediate consequence of theorem 3. Thus if axiom VII is assumed we can characterize the ordinals as those transitive sets whose elements are transitive. The ordinals were defined in this way by Gödel in his lectures at Vienna in 1937.

That this characterization of the ordinals is no longer sufficient if the axiom VII is dropped can be seen by considering again the possibility of a set which is its own only element.

3) Using V a, we can prove that a transitive set, every transitive proper subset of which is an element of it, is an ordinal. For let n be a transitive set, and suppose that every transitive proper subset of n is an element of n. By V a, the class of ordinals in n is represented by a set s; and, by theorem 3 and the transitivity of n, this set s is transitive. Thus s is a transitive set of ordinals and therefore an ordinal. This ordinal must be higher than every ordinal in n and consequently cannot be an element of n. Since every transitive proper subset of n is an element of n, it follows that s cannot be a proper subset of n. Therefore $s=n$, and n is an ordinal.

Since the converse of this follows from theorem 1, the ordinals can be characterized as those transitive sets having the property that every transitive proper subset is an element. We could take this as the definition of an ordinal and, using V a, could develop the theory of ordinals from this definition almost as easily as from the definition we have used. The main difference would be that instead of our theorem 1 we would have to prove that an ordinal cannot have a reflexive element; this we could do by showing that, for every ordinal n, the class of those elements of n which do not have a reflexive element is represented by n itself.

4) The definition of an ordinal on which Zermelo based his independent theory of ordinals, referred to at the beginning of this section, was stated in terms of his axioms for set theory (which have no distinction between classes and sets). It can, however, also be formulated with reference to our present system, as follows:[22]

A set n is an ordinal if (1) either $0=n$ or $0 \epsilon n$, and (2) if $a \epsilon n$ then $a'=n$ or $a' \epsilon n$, and (3) if s is a subset of n, the sum of the elements of s is represented by n or by an element n.

On the basis of the axioms I–III and V a, this definition can be proved equivalent to our (second) definition. Proofs of the theorems of the Zermelo independent theory of ordinals require in our system no other axioms than I–III and V a.

5) From the axioms I–III and VII it is not possible to prove that a sufficient condition for a transitive set to be an ordinal is that every transitive proper subset of it be an element of it. Likewise it is impossible to prove from these axioms that every set which is an ordinal according to Zermelo's definition, as just given, is also an ordinal under our definition. Both of these independence assertions will be established later by means of a model, or independence example. This model—like the one announced above, concerning the possibility of a set's being its own only element—will be on the basis of number theory as derived from our axioms I–III and VII.

[22] We here modify Zermelo's definition slightly, replacing his condition "$0 \epsilon n$" by "$0=n$ or $0 \epsilon n$," in order that the null set may be counted amongst the ordinals, as it is under our other definitions.

In the applications of ordinals which we shall make, it does not matter which particular definition of an ordinal is chosen. We need only the following properties:

0 is an ordinal.

If a is an ordinal, a' is an ordinal.

Every transitive set of ordinals is an ordinal.

The ordinals can be ordered by a relation "lower than" in such a way that the elements of an ordinal are all the ordinals lower than it.

To every non-empty class of ordinals belongs a lowest ordinal.

The class of all ordinals exists.

6. Number theory. From the fundamental theorems about ordinals we can pass to number theory by means of the notion of a *finite ordinal.*

Definition. An ordinal is *finite* if both it and every ordinal lower than it are either 0 or of the form c'; or, equivalently, if it and every element of it satisfy the condition of either having a highest element or having no element.

From this definition it follows at once that an ordinal lower than a finite ordinal is finite; also, by the class theorem, that the class of all finite ordinals exists. The following consequences are also immediate:

0 is a finite ordinal.

If n is a finite ordinal, then n' is a finite ordinal and $n' \neq 0$.

If m and n are finite ordinals and $m' = n'$, then $m = n$.

If C is a class to which the null set belongs, and if, for every finite ordinal n belonging to C, n' also belongs to C, then every finite ordinal belongs to C. (*Principle of complete induction.*)

In order to prove the principle of complete induction, we may reason as follows. If there were a finite ordinal which did not belong to C, the class of finite ordinals not belonging to C (which exists, by the class theorem) would have a lowest element; but this lowest element could be neither the null set nor of the form c'; hence a contradiction.

Thus the Peano axioms are satisfied for the finite ordinals if the null set is taken as the number 0 and n' as the successor of n.

Moreover the usual method of introducing functions of natural numbers by "recursive definitions" can be justified by an existence theorem.[23] For this purpose we first prove the following *iteration theorem*:

Let a be a set, A a class to which a belongs, and F a function which assigns to every element of A an element of A. Then there exists a function H (depending on the parameters a and F) which assigns to every finite ordinal an element of A, and satisfies the conditions,

$$H(0) = a,$$
$$H(n') = F(H(n))$$

for every finite ordinal n.

[23] Justification of recursive definition of numerical functions, from the set-theoretic point of view, was first given by Dedekind in ***Was sind und was sollen die Zahlen?***, Braunschweig 1888, §9. We here follow his method, but with the modification that we reduce the recursive definition to the more special case of an iteration.

We show first that, under the given hypotheses, there exists for every finite ordinal n a unique functional set s satisfying the conditions:

(1) The domain of the function represented by s is the class of all ordinals lower than or equal to n.

(2) $\langle 0, a\rangle \epsilon s$.

(3) If $\langle k, c\rangle \epsilon s$ and $k \epsilon n$, then $\langle k', F(c)\rangle \epsilon s$.

This we do as follows. The assertion about n that it is a finite ordinal and that there exists a unique functional set satisfying (for this n) the conditions (1), (2), (3) is a predicate of n definable by a constitutive expression, with a and F as parameters. Hence, by the class theorem, there exists the class C of those finite ordinals for which there exists a unique functional set s satisfying (1), (2), (3). Moreover we show easily that 0 belongs to C, and if a finite ordinal n belongs to C then n' belongs to C. Hence by complete induction we infer that C is the class of all finite ordinals.

Now the predicate of n and b, that n is a finite ordinal and b is the value assigned to n by a functional set satisfying (1), (2), (3), can be defined by a constitutive expression. Hence, by the class theorem, there exists the class of pairs $\langle n, b\rangle$ such that this predicate holds of n and b. It then follows that this class of pairs is a function H, whose domain is the class of all finite ordinals, and whose value $H(n)$ for a finite ordinal n is the value assigned to n by the functional set s satisfying (1), (2), (3). The equation $H(0)=a$ obviously holds. That $H(n')=F(H(n))$ for every finite ordinal n can be proved by showing—as is easily done—that the value, assigned to n by the functional set satisfying (1), (2), (3) with n' substituted for n, is the same as that assigned to n by the functional set s satisfying (1), (2), (3).

Finally, by an application of complete induction, we see that for every finite ordinal n the value of H belongs to A. This completes the proof of the iteration theorem.

Definition. If a is a set and F a function, and there exists a class A such that $a \eta A$ and, if $c \eta A$, then $F(c)\ \eta\ A$, then the function H, which has the class of finite ordinals as its domain and satisfies the conditions $H(0)=a$ and $H(n') = F(H(n))$ for every finite ordinal n, will be called the *iterator* of F on a.

The constitutive expression by which, under the hypotheses of the iteration theorem, the iterator of F on a is defined, contains a as a parameter, and the variable n ranging over the finite ordinals as an argument. But we may also take a as argument and n as parameter; or n and a may both be taken as arguments. Hence we obtain the following:

Corollary. On the hypotheses of the iteration theorem, there exists for every finite ordinal n a function assigning to every element a of A the value that the iterator of F on a has for n; this function will be called the *n-fold iteration* of F. Likewise there exists a function assigning to every pair $\langle a, n\rangle$, with $a \eta A$ and n a finite ordinal, the value for n of the iterator of F on a.

By means of the concept of the n-fold iteration of a function, the most familiar arithmetic functions can be defined. Thus for finite ordinals m, n:

$m+n$ is the value for the argument m of the n-fold iteration of the function which assigns to a finite ordinal a the value a'.

$m \cdot n$ is the value for the argument 0 of the n-fold iteration of the function which assigns to a finite ordinal a the value $a+m$.

m^n is the value for the argument $0'$ of the function which assigns to a finite ordinal a the value $a \cdot m$.

Under these definitions, we have for all finite ordinals m, n the equations,

$$m + 0 = m \qquad m \cdot 0 = 0 \qquad m^0 = 0'$$
$$m + n' = (m + n)' \qquad m \cdot n' = m \cdot n + m \qquad m^{n'} = m^n \cdot m$$

—from which the well-known properties of the three arithmetic functions can be derived by complete induction.

It is also possible to show that the ternary relations,

$$a + b = c, \qquad a \cdot b = c, \qquad a^b = c,$$

can be represented by constitutive expressions.

From the iteration theorem we deduce the following *theorem of finite recursion*: Let a be a set, A a class to which a belongs, and G a function which assigns an element of A to every pair $\langle k, c\rangle$ in which k is a finite ordinal and c is an element of A. Then there exists a function K which assigns an element of A to every finite ordinal, and satisfies the conditions,

$$K(0) = a,$$
$$K(n') = G(\langle n, K(n)\rangle)$$

for every finite ordinal n.

For let P be the class of all pairs $\langle k, c\rangle$, where k is a finite ordinal and $c\eta A$, and let F be the function obtained from G by replacing every element $\langle\langle k, c\rangle, d\rangle$ of G by $\langle\langle k, c\rangle, \langle k', d\rangle\rangle$. (The existence of such a function follows by means of the axioms III from the existence of the class of pairs $\langle k, k'\rangle$, which in turn follows from the class theorem.) Then $\langle 0, a\rangle\eta P$, and F assigns to every element of P an element of P. Now, using the iteration theorem, let H be the iterator of F on $\langle 0, a\rangle$. Then,

$$H(0) = \langle 0, a\rangle,$$
$$H(n') = F(H(n))$$

for every finite ordinal n. It follows by complete induction that, for every finite ordinal n, $H(n)$ is a pair having n as its first member. Hence

$$F(H(n)) = \langle n', G(H(n))\rangle.$$

If K is the converse domain of H, it follows that K is a function such that, for every finite ordinal n,

$$H(n) = \langle n, K(n)\rangle.$$

Then

$$K(0) = a,$$

and, for every finite ordinal n,

$$\langle n', K(n')\rangle = \langle n', G(H(n))\rangle,$$
$$K(n') = G(H(n))$$
$$= G(\langle n, K(n)\rangle).$$

This completes the proof.

Remark 1. The theorem of finite recursion obviously could be proved directly in the same way that we proved the iteration theorem (and the iteration theorem itself would then be an immediate corollary). However, it seems to be of interest in itself that a "recursive definition" of the form $K(0)=a$, $K(n')=G(\langle n, K(n)\rangle)$ can be reduced to a definition by mere iteration of a function.[24]

Remark 2. As can be seen by examining the proof, there is a *corollary* to the theorem of finite recursion entirely analogous to that which we stated for the iteration theorem.

We have also the means of reproducing definitions of numerical functions employing the concept of "the least number such that." This results from the following considerations.

Let C be a class of pairs whose domain is a class of ordinals and converse domain a subclass of a class A. Consider the following predicate of a and b: "a belongs to A; if a does not belong to the converse domain of C, b is the null set; and if a does belong to the converse domain of C, then b is the lowest of those ordinals which occur as first member in some pair which belongs to C and has a as second member." This predicate can be defined by a constitutive expression. Hence, by the class theorem, there exists a function having A as its domain and assigning (1) to each element a of A for which there is an ordinal n such that $\langle n, a\rangle \eta C$, the lowest such ordinal, and (2) to other elements of a, the null set.

This applies in particular to the case that, for a given $(k+1)$-ary relation $R(a_1, a_2, \cdots, a_k, a_{k+1})$ between finite ordinals which can be represented by a constitutive expression, C is the class (whose existence follows from the class theorem) of those normal $(k+1)$-tuplets $\langle a_{k+1}, \langle a_1, \langle a_2, \cdots, \langle a_{k-1}, a_k\rangle \cdots\rangle\rangle\rangle$, formed out of finite ordinals, for which $R(a_1, a_2, \cdots, a_k, a_{k+1})$ holds. Here the elements of the domain of C are finite ordinals, and the elements of the converse domain are normal k-tuplets formed out of finite ordinals.

Thus we obtain the result that, for every $(k+1)$-ary relation $R(a_1, a_2, \cdots, a_k, a_{k+1})$ between finite ordinals which can be represented by a constitutive expression, there exists a function which assigns to every normal k-tuplet $\langle a_1, \langle a_2, \cdots, \langle a_{k-1}, a_k\rangle \cdots\rangle\rangle$ whose members are finite ordinals (1) the lowest finite ordinal a_{k+1} such that $R(a_1, a_2, \cdots, a_k, a_{k+1})$ holds, if there is such a finite ordinal a_{k+1} at all, and (2) otherwise the ordinal 0.

With this, we have at our disposition a sufficient basis for the development of number theory.[25]

We shall now indicate how the *theory of finite classes and finite sets* is obtainable from the theory of finite ordinals. For this purpose we require first the following results about finite ordinals.

[24] This reduction figures in the investigations of A. Church, J. B. Rosser, and S. C. Kleene on the "λ-definability" of numerical functions. See S. C. Kleene, *A theory of positive integers in formal logic*, **American journal of mathematics,** vol. 57 (1935), pp. 153–173, 219–244, especially the passage from the theorem 15 IV to 15 V, pp. 220–221.

[25] Cf., e.g., Hilbert and Bernays, ***Grundlagen der Mathematik,*** vol. 1, pp. 401–422.

(α) *If the domain of a function is represented by a finite ordinal, the converse domain is represented by a set.*

For let F be the function in question and n the finite ordinal representing the domain of F. By complete induction we can show that for every finite ordinal k there exists a set whose elements are those sets c for which there exists an element $\langle l, c\rangle$ of F with $l \epsilon k$. Then taking n for k we have that the converse domain of F is represented by a set.

Corollary 1. Every subclass of a finite ordinal is represented by a set.

For if S is a subclass of a finite ordinal n, then either S is empty and is represented by the null set, or for any element a of S the class whose elements are the pairs $\langle c, c\rangle$ for which $c \eta S$ and the pairs $\langle d, a\rangle$ for which $d \epsilon n$ but not $d \eta S$, is a function whose domain is represented by n and whose converse domain is S. Hence, by (α), S is represented by a set.

Corollary 2. If there is a one-to-one correspondence between a class A and a finite ordinal, then A and the one-to-one correspondence are each represented by a set.

This follows by applying (α) to the converse class of the one-to-one correspondence in question, and to the class of triplets of the form $\langle a, \langle b, a\rangle\rangle$ where $\langle b, a\rangle$ belongs to the one-to-one correspondence—since each of these classes is a function whose domain is a finite ordinal.

Definition. That there exists a one-to-one correspondence between the class A and the class B will be denoted by $A \backsim B$ ("A one-to-one with B").

This relation has the characteristic properties of an equality or equivalence relation. In fact, for every class A we have $A \backsim A$ by consequence 6 of the axioms III, and the axioms III b (3), a (3); if $A \backsim B$ then $B \backsim A$, by III c (2); if $A \backsim B$ and $B \backsim C$, then $A \backsim C$, by the composition lemma.

For sets a and b, a representing a class A, and b representing a class B, the notations $a \backsim b$, $a \backsim B$, $A \backsim b$ will be used to mean the same thing as $A \backsim B$. (Read "a one-to-one with b," etc.)

Corollary 3. There exists the class of those pairs $\langle a, b\rangle$ for which a is a finite ordinal and $a \backsim b$.

This follows from corollary 2 by the same method by which in §4 the existence of the class of all pairs $\langle a, b\rangle$ for which $a \backsim b$ was inferred[26] from the consequence 3 of axiom V b. (For the present corollary, of course V b is not required.)

(β) *For every proper subset s of a finite ordinal n, there exists a one-to-one correspondence between s and an ordinal lower than n.*

This is proved as follows. By the class theorem and corollary 3 of (α) there exists a class A whose elements are those finite ordinals n which satisfy the condition that, for every proper subset s of n, there is a one-to-one correspondence between s and an ordinal lower than n. Evidently 0 belongs to A. Thus, in order to prove (β), by the principle of complete induction it is sufficient to show that, if a finite ordinal n belongs to A, then n' belongs to A.

Let n be a finite ordinal belonging to A and t a proper subset of n'. If $t = n$, then $t \backsim n$, and so there is a one-to-one correspondence between t and an ordinal lower than n'; and this holds also for $t \subset n$, since $n \eta A$. If t is not a subset of n,

[26] See p. 2 above.

then $n\epsilon t$ and there is at least one element of n that is not in t. Let P be the class of those elements of t which are different from n. By corollary 1 of (α), P is represented by a set p. Since $p \subset n$ and $n \eta A$, there exists a one-to-one correspondence C between p and an ordinal m lower than n. Now the class, whose elements are the elements of C and in addition the pair $\langle n, m\rangle$, is a one-to-one correspondence between t and m'; and m' is an ordinal lower than n'. Thus in every case there exists a one-to-one correspondence between t and an ordinal lower than n'—as was to be proved.

(γ) There cannot be a one-to-one correspondence between a finite ordinal n and an ordinal different from n.

The proof is as follows. By the class theorem and corollary 3 of (α), there exists the class B of those finite ordinals of which there exists a one-to-one correspondence to another ordinal. If this class were non-empty, there would be a lowest ordinal l belonging to it and there would exist an ordinal k different from l such that $l \sim k$. This ordinal k could not be lower than l, since otherwise we would have $k \eta B$ in contradiction to the characterizing property of l. Hence $l \subset k$; and consequently the class of those pairs, belonging to the one-to-one correspondence between l and k, which have their second member in l would be a one-to-one correspondence between a proper subclass of l and l itself. That subclass being represented, in accordance with corollary 1 of (α), by a set m, we should have $m \subset l$ and $m \sim l$. Moreover, by (β), there would exist an ordinal p lower than l such that $m \sim p$, and from $m \sim l$ and $m \sim p$ would follow $p \sim l$, and hence $p \eta B$, in contradiction to the characterizing property of l. It follows that the class B is empty, as was to be proved.

Now we are able to go on to the general definition of finiteness.

Definition. A class or set will be called *finite* if there exists a one-to-one correspondence between it and a finite ordinal.

In order to show that this definition is really a generalization of our definition of a finite ordinal, we have to verify that according to it an ordinal is a finite set if and only if it is a finite ordinal. This is now immediate, however, since, on the one hand, every finite ordinal is one-to-one with itself, and, on the other hand, by (γ), there cannot be a one-to-one correspondence between an ordinal that is not finite and a finite ordinal.

We note also the existence of the class of all finite sets—which follows from corollary 3 of (α).

As consequences of (α), (β), (γ), we obtain further the following *theorems on finite classes and sets.*

1. Every finite class is represented by a finite set.—This is an immediate consequence of corollary 2 of (α).

2. Every subclass of a finite class is finite.—This follows by combining corollary 1 of (α) with (β).

3. There is only one finite ordinal which is one-to-one with a given finite class (or finite set).—This is an immediate consequence of (γ).

Definition. We shall call the unique finite ordinal which is one-to-one with a given finite class A (or a given finite set a) the *number attributable* to the class A (or to the set a).

4. The number attributable to a proper subclass (or a proper subset) of a

finite class (or a finite set) is lower than the number attributable to this class (set).—We obtain this as a consequence of 3 (just stated) and (β), making use of the fact that every subclass of a finite ordinal is represented by a set.

5. There is no one-to-one correspondence between a finite class and a proper subclass of it.

For let A be a finite class, B a proper subclass of it, and n the number attributable to A. Then n is a finite ordinal and there exists a one-to-one correspondence C between A and n. The class of those elements of C whose first member belongs to B is a one-to-one correspondence between B and a proper subclass of n that is represented by a set s. By (β), there exists an ordinal m lower than n such that $s \backsim m$. Now if there were a one-to-one correspondence between A and B, we should have $A \backsim B$, $n \backsim A$, $B \backsim s$, $s \backsim m$, and consequently $n \backsim m$. But this contradicts (γ).

Starting from these fundamentals, the theory of finite classes and sets and of the numbers attributable to them can be derived by number-theoretic methods. In particular, the following theorems are to be proved here:

1) If A and B are finite classes and have no common element, and m and n are the numbers attributable to A and B respectively, then the sum of A and B is finite and the number attributable to it is $m+n$.

2) If A and B are finite classes and m and n are the numbers attributable to them, then the class of pairs $\langle a, b\rangle$ such that $a\eta A$ and $b\eta B$ is finite and the number attributable to it is $m\cdot n$.

3) If A is a finite class and m is the number attributable to it, then the class of subsets of A is finite and the number attributable to it is 2^m, where 2 is the ordinal $0''$.

4) If A is a finite class whose elements are finite, then the sum of the elements of A is finite.

A principal point of the method here set forth for developing number theory and the theory of finite classes and sets is that it is necessary to use for it, besides the axioms I–III, only the restrictive axiom VII or instead of it the subclass axiom V a.

For the foundation of number theory by Dedekind's method—which can also be reproduced within our system—the axiom of infinity VI would be required, as well as V a. Zermelo was the first to show that, for a set-theoretic foundation of number theory and the theory of finite sets, it is possible to do without assuming the existence of an infinite set.[27]

Analogous to this possibility for number theory is the possibility of avoiding the sum axiom V c and the power axiom V d in the development of analysis and of general set theory—as will appear in the next two parts.

ZURICH

[27] E. Zermelo, *Sur les ensembles finis et le principe de l'induction complète*, ***Acta mathematica***, vol. 32 (1909), pp. 185–193 (dated 1907); *Ueber die Grundlagen der Arithmetik*, ***Atti del IV Congresso Internazionale dei Matematici (Roma, 6–11 aprile 1908)***, vol. 2 (1909), pp. 8–11. See also Kurt Grelling, ***Die Axiome der Arithmetik***, dissertation Göttingen 1910.